View metadata, citation and similar papers at core.ac.uk

brought to you by CORE

provided by Kanazawa University Repository for Academic Resources

Criterial image preparation for a search method for digital watermarking image using correlation coefficient between pixel value histograms

著者	Imamura Kousuke, Kuroda Hideo, Fujimura Makoto
journal or publication title	ICICS 2013 - Conference Guide of the 9th International Conference on Information, Communications and Signal Processing
page range	6782916
year	2013-12-01
URL	http://hdl.handle.net/2297/37542

doi: 10.1109/ICICS.2013.6782916

Criterial Image Preparation for a Search Method for Digital Watermarking Image Using Correlation Coefficient Between Pixel Value Histograms

Kousuke Imamura
Kanazawa University
Kakuma-machi, Kanazawa, Ishikawa
920-1192 Japan

Hideo Kuroda
FPT University
Cau Giay, Hanoi, Vietnam

Makoto Fujimura
Nagasaki University
1-14 Bunkyo-machi, Nagasaki
852-8521 Japan

***Abstract*— An efficient two-stage image search method for the extraction of illegal copies on the Internet has been proposed in our previous paper. In the present paper, we propose a criterial image preparation method for more efficient searching in the pre-search stage. It is possible to accelerate searching without omission because the criterial image preparation method provides a criterial image set for the various attacks involving changes to known pixel values. In addition, we propose an improved pre-search algorithm that includes the extraction of a GIF-compression-attacked image involving color number transformation. The performance of the improved pre-search method is evaluated through computer simulation.**

Keywords- digital watermark, image search, criterial image preparation, correlation coefficient, pixel value histogram

I. INTRODUCTION

Multimedia digital content has been used in various fields due to the development of computer and network technology. Moreover, the generation and transcription of digital content including images and the sending and receiving of large amounts of content over the Internet became sufficiently easy that it can be done by individuals. However, although convenient, this has also led to copyright infringement due to unauthorized copying of content by third parties.

Digital watermarking is one method for the prevention of copyright infringement. The digital watermarking technique actually consists of two methods: embedding copyright information in content and extracting particular content from a larger amount of content. Although various digital watermarking techniques have been investigated [1]-[3], these studies discussed only methods for the embedding and extraction of copyright information into/from content, whereas a method for effectively searching for illegal copies in a large amount of content on the Internet has not yet been sufficiently considered.

Image searching techniques using feature vectors representing the characteristics of particular images have been proposed [4][5]. However, these methods cannot be applied to illegal copies in the large amounts of image content on the Internet, because the images on the Internet are not connected with the feature vector.

On the other hand, reverse image search techniques [6][7] can be applied to the large amounts of image content on the Internet. In such techniques, a feature vector of owned image content is extracted in advance, and the feature vector is registered at a certificate center. If the extracted feature vector from an image on the Internet matches the feature vector of the particular registered image at the certificate center, then an owned image has been identified. In applying such methods to an image search for a digital watermark, it is possible to use the copyright information embedded by a digital watermark as the feature vector for the image search. However, the above example does not allow high efficiency because the procedure for extracting embedded information from an image is generally highly complex.

In order to solve the above problem, we have proposed an efficient two-stage image search method for the extraction of illegal copies of a target image [8]. The pre-search algorithm of the proposed method introduced a correlation coefficient between pixel value histograms of images for the image matching decision, and the algorithm searched for illegally copied images, including small changes in pixel values. The image extracted by the pre-search algorithm was a geometrically attacked image, for example by rotation and/or scaling, or a JPEG-compression-attacked image. In the extraction of the JPEG-compression-attacked image, JPEG-compressed images with quality parameters of 60 and 20 were defined as a criterial image set which were the reference images for the search. However, since the amount of change of the pixel values in JPEG-compressed images depends on the features of the images, there is no guarantee that extraction using the two criterial images will provide a complete search of JPEG-compressed images.

With this in mind, we first propose a criterial image preparation method for the pre-search algorithm in the present paper. The proposed criterial image preparation method selects a criterial image based on extraction verification of the attacked target images and can search efficiently without extractive omission. Second, we add a criterial image set obtained using

the proposed criterial image preparation method for the extraction of GIF-compression attack with color transformation. In the present paper, we improve the pre-search algorithm by introducing the above proposal, and the performance of the improved pre-search algorithm is evaluated through computer simulations.

II. TWO-STAGE SEARCH METHOD FOR DIGITAL WATERMARKING IMAGES

We have proposed a two-stage image search method in our previous paper [8]. The two-stage image search method consists of pre-search and main search stages. In the pre-search stage, image content from the Internet is input and whether this content is similar to owned content is investigated using a simple method. If the content is similar to owned image content, the image content is output to the main search stage as an illegal copy candidate. If the content is not similar, the search process is terminated and the next image content is investigated. In the main search stage, the embedded copyright information in the image is extracted by a predefined method. When the extracted copyright information indicates an owned image, a complaint is sent to the user through a predefined procedure.

A. Image Search Using Correlation Coefficient between Pixel Value Histograms

In the pre-search of the two-stage image search, it is necessary to extract candidate illegal copies by a simple method in order to achieve efficient extraction from among the large number of images on the Internet. In addition, robustness to small changes in pixel values is required in order to avoid omissions because the pixel values of illegal copies may include various noises used to attack the watermarking. With this in mind, pixel value histograms are used as feature quantities for image searches because histograms are insensitive to attack by geometric transform, e.g., rotation and scaling.

The correlation coefficient is defined as follows:

$$r = \frac{\sum_{i=0}^{255}(x_i - \bar{x})(y_i - \bar{y})}{\sqrt{\sum_{i=0}^{255}(x_i - \bar{x})^2 \sum_{i=0}^{255}(y_i - \bar{y})^2}} \qquad (1)$$

where $x(=\{x_i\})$, $y(=\{y_i\})$ denote the normalized occurrence frequencies at pixel value i, and $\bar{x}$ and $\bar{y}$ denote arithmetic means.

B. Pre-search Algorithm

The pre-search algorithm in the previous version extracts a candidate illegally copied image using the original image and JPEG-compressed images with quality parameters 60 and 20. In this algorithm, an image on the Internet is input, and the correlation coefficient r_o between the pixel value histograms of the original image and the observed image is calculated. If r_o is greater than or equal to Th_o, the observed image is extracted as a candidate illegally copied image. If r_o is less than Th_o, the correlation coefficients r_{q60} and r_{q20} between the smoothed histograms of the JPEG-compressed images with quality parameters 60 and 20, respectively, of the observed image are calculated. If r_{q60} is greater than Th_{q60} or if r_{q20} is greater than Th_{q20}, the observed image is extracted as a candidate. Otherwise, the observed image is excluded from the input of the main search.

III. CRITERIAL IMAGE PREPARATION METHOD IN PRE-SEARCH

For example, the JPEG image compression format has a quality parameter of from 1 to 100 for decision of the reconstructed image quality and compression rate. The reconstructed image quality gradually deteriorates so that the quality parameter becomes small, and the pixel value of the reconstructed image changes greatly from that of the original image. There is no guarantee that extraction using the two criterial images will provide a complete search of JPEG-compressed images, because the amount of change of pixel values in JPEG-compressed images depends on the features of the image. Conversely, because JPEG-compressed images with all quality parameters are sometimes extracted by only using original image, the pre-search method can be improved. In order to solve these problems, we propose a criterial image preparation method for pre-search. The criterial image preparation method selects criterial images based on extraction verification of target attacked images, and efficient searching is possible without omission.

A. Criterial Image Preparation for JPEG-Compression Attack

We describe the proposed criterial image preparation algorithm for the extraction of JPEG-compression-attacked images with all quality parameters. Fig. 1 shows the criterial image preparation algorithm for JPEG-compression-attack.

First, a JPEG-compression-attacked image set is constructed by compressing an original image using the JPEG compression format with quality parameters of from q_{min} to 100. Here, q_{min} is a configurable parameter because the deterioration of JPEG-compressed images with very low quality parameters is significant, and the utility value as illegal copies is poor. In the present paper, the parameter q_{min} is set to 20.

Second, the correlation coefficient $r_{o.}$ between the pixel value histogram of the original image and the JPEG-compressed images are calculated. A JPEG-compressed image having a correlation coefficient r_o that is greater than Th_o is extracted as a target image. For the case in which all of the JPEG-compressed images are extracted, the criterial image for a JPEG-compression attack is only an original image. If all of the JPEG-compressed images are not extracted, the unextracted JPEG-compressed image having minimal r_o is added to the criterial image set.

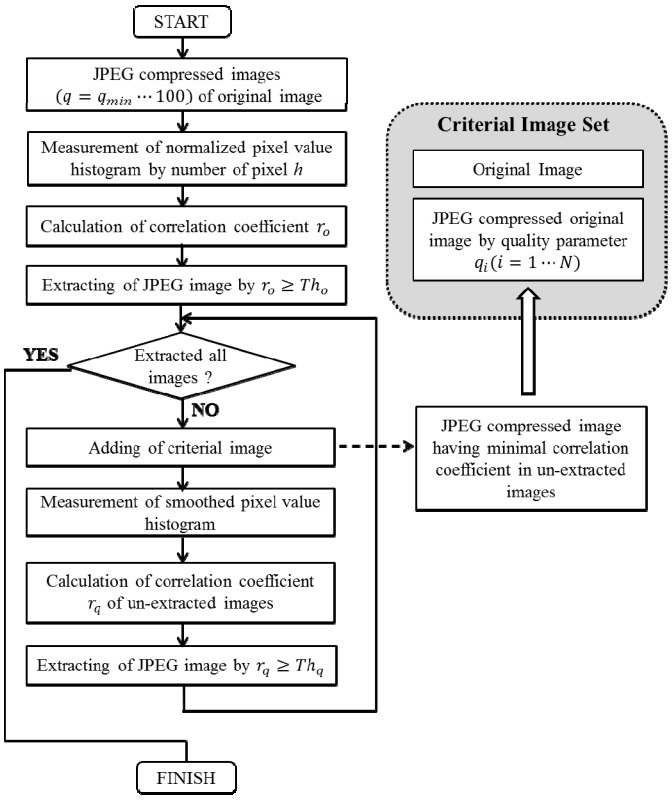

Figure 1. Criterial image preparation algorithm for JPEG-compressed attack

Next, the correlation coefficient r_q between the five- level smoothed pixel value histograms of the added criterial image and the unextracted JPEG-compressed images is calculated. If r_q is greater than or equal to Th_q, the observed image is extracted as a JPEG-compressed target image. If the all of the JPEG-compressed images are not extracted, the previous estimated JPEG-compressed image having minimal r_q is added to the criterial image set. The criterial image addition and selection process is iterated until all of the JPEG-compressed images are extracted. The images obtained by the above process become the criterial image set for the JPEG-compression attack.

B. Criterial Image Preparation for GIF-Compression Attack

Many images on the Internet are compressed by image coding techniques in order to enable efficient transmission. There are a number of compression formats besides the JPEG compression format, which is the de facto standard. Commonly known image compression formats include GIF and PNG. Based on the above considerations, it is necessary to design the pre-search algorithm so that its searches include these image compression formats. GIF is a lossless compression format that uses the Lempel-Ziv-Welch (LZW) data compression algorithm. The maximum number of colors handled by GIF is 256 and the number of colors must be a power of two. Therefore, GIF image compression includes a subtractive color process that has a significant effect on the pixel value histogram of an image. Moreover, countermeasures against subtractive color attack are also necessary.

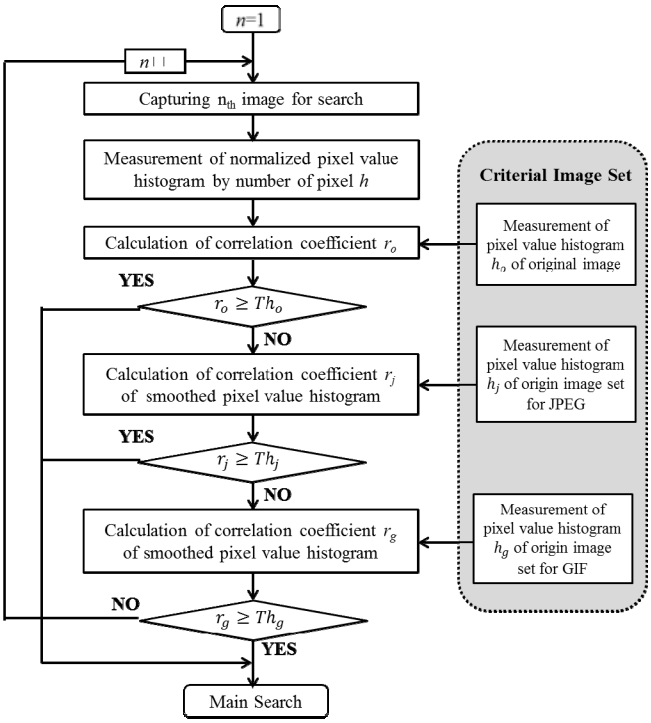

Figure 2. Pre-seaerch algorithm

In the present paper, the criterial images for GIF-compression attack are selected by the proposed criterial image preparation method, and the obtained criterial images are added to the criterial image set in pre-search. On the other hand, PNG supports lossless and full-color image compression. Therefore, PNG-compression-attacked images are extracted using the original image.

IV. PRE-SEARCH ALGORITHM IN THE PROPOSED IMAGE SEARCH METHOD

Fig. 2 shows the improved pre-search algorithm in the proposed image search method. In the pre-search algorithm, the correlation coefficient r_o between pixel value histograms of the original image and the observed image is first calculated. If r_o is greater than or equal to Th_o, the observed image is extracted as a candidate illegally copied image. Up to this point, the algorithm is effective for unattacked images and geometrically attacked images.

If r_o is less than Th_o, the correlation coefficients r_j between the smoothed histograms of the criterial image set for a JPEG-compressed attack and the observed image are calculated. If r_j is greater than or equal to Th_j, the observed image is extracted as a candidate illegally copied image. If r_j is less than Th_j, the correlation coefficients r_g between the smoothed histograms of the criterial image set for a GIF-compressed attack and the observed image are calculated. If r_g is greater than or equal to Th_g, the observed image is extracted as a candidate. If r_g is less than Th_g, the observed image is excluded from the input of the main search. For the case in which the observed image is

extracted as a candidate illegally copied image by the process to this point, the obtained candidate is output to the main search. In the present paper, Th_o, Th_j and Th_g are set to 0.95.

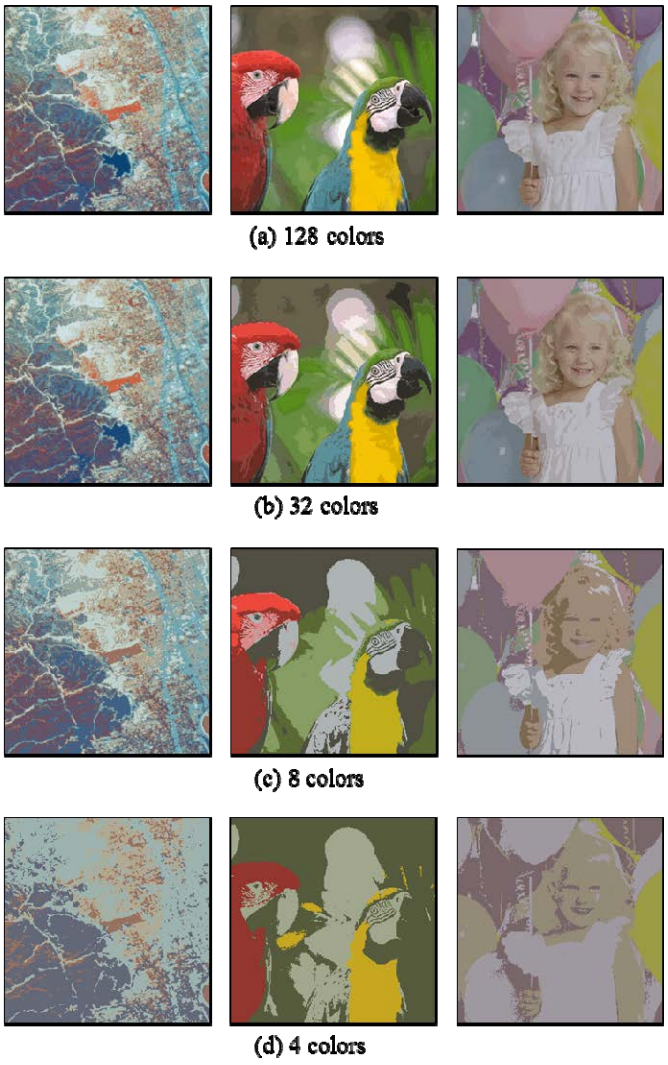

Figure 3. Color quantization of images

V. SIMULATION AND RESULTS

We analyzed the improved pre-search algorithm in the proposed method by computer simulations. In the experiment, 12 images (256x256 pixels, 24-bit color) from SIDBA, i.e., "Aerial", "Airplane", "Balloon", "Earth", "Girl", "Lenna", "Mandrill", "Parrots", "Pepper", "Sailboat", "Couple", and "Milkdrop", are used as test images.

A. Criterial Image for JPEG-Compression Attack

Firstly, we verify the criterial image selected by the proposed criterial image preparation method using JPEG compressed images. Table 1 shows the number and quality parameter of criterial images for JPEG-compression attack selected by the proposed criterial image preparation method. The quality parameters of the target JPEG-compression image are set to from 20 to 100.

TABLE I. NUMBER OF ADDED CRITERIAL IMAGES AND QUALITY PARAMETERS FOR A JPEG-COMPRESSION ATTACK

Image	Criterial image for JPEG-Compression	
	Number of added criterial images	***Quality parameter (Number of extracted image)***
Aerial	0	
Airplane	1	22(22)
Balloon	1	25(15)
Earth	4	21(5),20(6),38(32),24(8)
Girl	1	23(16)
Lenna	1	20(13)
Mandrill	0	
Parrots	1	23(18)
Pepper	1	20(7)
Sailboat	1	20(17)
Couple	1	21(41)
Milkdrop	1	20(18)

TABLE II. NUMBER OF ADDED CRITERIAL IMAGES AND COLOR FOR A GIF-COMPRESSION ATTACK

Image	Criterial image for GIF-Compression	
	Number of added criterial images	***Number of Color (Number of extracted image)***
Aerial	7	4(1), 8(1), 64(1), 16(1), 128(1), 32(1), 256(1)
Airplane	7	
Balloon	7	
Earth	7	
Girl	7	
Lenna	7	
Mandrill	7	
Parrots	7	
Pepper	7	
Sailboat	7	
Couple	7	
Milkdrop	7	

A value of 0 for the number of added criterial images indicates that the criterial image for JPEG-compression attack is not necessary because the all of the JPEG-compression-attacked images are extracted using the original image. Table 1 shows that adding one criterial image to the criterial image set for a JPEG-compression attack is often sufficient. As a result, the average number of the criterial images is fixed to less than two in the previous method. Therefore, the criterial image set selected by the proposed method will provide faster searching without omission of extraction.

B. Criterial Image Preparation for GIF-Compression Attack (Color Number Transformation)

Next, we verify the criterial image for extracting GIF-compression-attacked images with subtractive color transformation. The number of supporting colors in GIF is a power of two. Fig. 3 shows the example of the color quantized images. In the present paper, since images with too few colors

image are not useful for illegal copied content, the minimum number of color is set to 4.

Table 2 shows the number of criterial images selected by the proposed criterial image preparation algorithm and the number of color quantization in GIF-compression attacks. From Table 2, the number of added criterial images is seven in all images, and a criterial image is required for each color-quantized image. This is because the pixel value histogram is greatly changed by the color quantization. As a result, we add seven criterial images for the extraction of a GIF-compression-attacked image.

C. Extracting Estimation of Pre-search Algorithm

We extracted JPEG-compression-attacked images and GIF-compression-attacked images. In the improved pre-search algorithm, the original image, the JPEG-compression-attack image, and the GIF-compression-attacked image are guaranteed to be extracted from all test images. However, if frequent excessive extraction, in which different images are extracted as the same image, occurs, then the extraction performance will decrease. Thus, we evaluate excessive extraction for all combinations of attacked images. As a result, excessive extraction did not occur.

VI. CONCLUSIONS

In the present paper, we proposed a criterial image preparation method for the pre-search algorithm of the efficient two-stage image search method. In addition, we incorporated the extraction of a GIF compression attacked image in the pre-search algorithm. Simulation results revealed that the pre-search method including the proposed criterial image preparation provides better performance in an image search for a digital watermark. In the future, we intend to develop a search technique for other type of attack on digital watermarks in images.

REFERENCES

[1] C.-W. Tang and H.-M. Hang, “A Feature-Based Robust Digital Image Watermarking Scheme,” IEEE Transactions on Signal Processing, Vol.51, No.4, pp.950-959, 2003.

[2] C.-Y. Lin et al., “Rotation, Scale, and Translation Resilient Watermarking for Images,” IEEE Transactions on Image Processing, Vol.10, No.5, pp.767-792, 2001.

[3] I.J. Cox, M.L. Miller, J.A. Bloom, J. Fridrich and T. Kalker, “Digital Watermarking and Steganography,” Morgan Kaufmann, 2008.

[4] T. Nocke and H. Schmann, “Meta Data for Visual Data Mining,” Proceedings of Computer Graphics and Imaging (CGIM’02), 2002.

[5] M.B.H. Hmida and Y. Slimani, “Meta-learning in Grid-based Data Mining Systems,” International Journal of Computer Network and Communications (IJCNC), Vol.2, No.5, pp.215-230, 2010.

[6] Hak-Soo Ju et al, “An Anonymous Buyer-Seller Watermarking Protocol with Anonymity Control,” Springer, Information Security and Cryptology, ICISC2002, pp.421-432, 2003.

[7] D. Birk and S. Gaines, “Using Digital Watermarking for Securing Next Generation Media Broadcasts,” Springer, Advances in Machine Learning and Data Analysis, pp.27-41, 2010.

[8] K. Imamura, H. Kuroda and M. Fujimura, “Image Content Detection Method Using Correlation Coefficient between Pixel Value Histograms,” Springer CCIS 260, pp.1-9, 2011.

[9] GIMP: The GNU Image Manipulation Program.